POLITECNICO DI TORINO
Repository ISTITUZIONALE

User-centric Networks Selection with Adaptive Data Compression for Smart Health

Original
User-centric Networks Selection with Adaptive Data Compression for Smart Health / Abdellatif, ALAA AWAD ABDELHADY; Mohamed, Amr; Chiasserini, Carla Fabiana. - In: IEEE SYSTEMS JOURNAL. - ISSN 1932-8184. - STAMPA. - 12:4(2018), pp. 3618-3628. [10.1109/JSYST.2017.2785302]

Availability:
This version is available at: 11583/2694525 since: 2018-12-06T14:59:07Z

Publisher:
IEEE

Published
DOI:10.1109/JSYST.2017.2785302

Terms of use:
openAccess

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2018 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

05 May 2023

User-centric Networks Selection with Adaptive Data Compression for Smart Health

Alaa Awad*,†, Amr Mohamed*, and Carla-Fabiana Chiasserini†
*Department of Computer Science and Engineering, Qatar University, Doha, Qatar
†Department of Electronics and Telecommunications, Politecnico di Torino, Italy
E-mail: {aawad, amrm}@qu.edu.qa and chiasserini@polito.it

Abstract

The increasing demand for intelligent and sustainable healthcare services has prompted the development of smart health systems. Rapid advances in wireless access technologies and in-network data reduction techniques can significantly assist in implementing such smart systems through providing seamless integration of heterogeneous wireless networks, medical devices, and ubiquitous access to data. Utilization of the spectrum across diverse radio technologies is expected to significantly enhance network capacity and Quality of Service (QoS) for emerging applications such as remote monitoring over mobile-health (m-health) systems. However, this imposes an essential need to develop innovative mechanisms networks selection that account for energy efficiency while meeting application quality requirements. In this context, this paper proposes an efficient networks selection mechanism with adaptive compression for improving medical data delivery over heterogeneous m-health systems. We consider different performance aspects, as well as networks characteristics and application requirements, so as to obtain an efficient solution that grasps the conflicting nature of the various users' objectives and addresses their inherent tradeoffs. The proposed methodology advocates a user-centric approach towards leveraging heterogeneous wireless networks to enhance the performance of m-health systems. Simulation results show that our solution significantly outperforms state-of-the-art techniques.

Index Terms

Heterogeneous wireless environment, network selection, multi-RAT architectures, decomposition.

I. Introduction

The increasing number of elderly individuals and chronic disease patients that require continuous monitoring within higher population cities, and the increasing demand for intelligent, efficient, and sustainable healthcare services, have led to the appearance of *mobile-health* concept. The evolution

This work was made possible by GSRA grant # GSRA2-1-0609-14026 and NPRP grant # 7-684-1-127 from the Qatar National Research Fund (a member of Qatar Foundation). The findings achieved herein are solely the responsibility of the authors.

of computational intelligence systems, mobile communications, and Internet of Things (IoT) has boosted the evolution of traditional healthcare processes into smart health services. M-health can be conceptualized as the integration of technologies (e.g., cloud computing, IoT, edge computing), and devices (e.g., smart bio-sensors and wearable devices) for providing reliable, and secure healthcare services (e.g., remote monitoring, smart hospitals, etc.). In general, smart health can be considered as the context-aware evolution of mobile-health (m-health), leveraging mobile technologies to provide smart personalized health through efficient and user-friendly services and tools that patients can use to monitor their own health conditions [1]. In such systems, a combination of implantable or wearable medical and non-medical sensors is leveraged for monitoring vital signs within the smart assisted living homes, which facilitates continuous monitoring and automatic detection of individuals' context [2].

Remote monitoring applications typically require the recording, processing, and transmission of large volumes of data. Consider, for instance, high-quality Electroencephalography (EEG) devices consisting of up to 100 electrodes that can generate a data rate of 1.6 Mbps per single patient. In normal conditions, such medical data should be reported to the M-Health Cloud (MHC) every 5 minutes, while, in the case of emergency where high-intensive monitoring is needed, it should be reported every 10 seconds [3]. This demand of high data rates and QoS has motivated us to leverage the development of cellular networks into dense heterogeneous networks (HetNets) with the utilization of multi-Radio Access Technology (RAT). However, it is essential for each user/device to leverage different RATs, hence, the available radio resources across different spectral bands, to communicate with the network infrastructure [4].

In such heterogeneous multi-radio environment, implementing m-health systems is challenging. As shown in Figure 1, we can divide m-health architecture into three main modules: data acquisition and pre-processing, wireless multi-RAT network, and health monitoring services and applications. Patients are equipped with sensor nodes, namely, a Body Area Sensor Network (BASN), pre-attached to their smartphones. The smartphone acts as Personal/Patient Data Aggregator (PDA), i.e., it gathers the medical data from the sensors and performs in-network processing to optimize the data transport based on the context and the network state. In particular, the PDA may compress the data at the cost of a certain degree of signal distortion, and sends it to the MHC over the multi-RAT wireless network. Importantly, the multi-RAT wireless environment allows the PDA to be connected anywhere and anytime, provided that innovative network selection techniques are implemented. Additionally, data transfer from the PDAs to the MHC should take place in an energy-efficient manner, in order to ensure a long lifetime of the battery-operated PDAs.

Accordingly, the focus of this paper is on how to benefit from the integration of multiple RATs

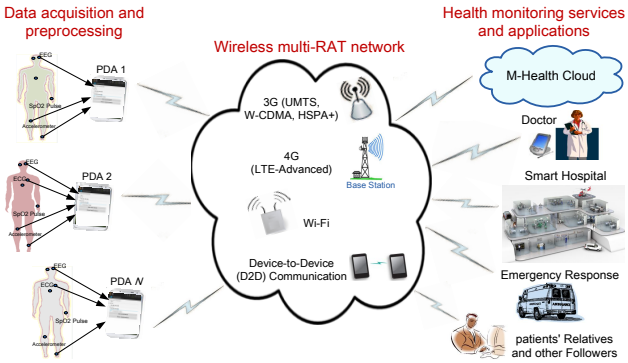

Fig. 1. Multi-RAT m-Health system scenario.

within the above system architecture, so as to meet the application and system requirements and to facilitate the management of, e.g., chronic diseases. Also, since PDAs can compress the data before transmission, each PDA has to select not only the most suitable RAT(s) through which data should be sent, but also the compression level to be used based on the network state. More specifically, we aim at answering the following questions: *(i)* Which network(s) should be selected among multiple RANs? *(ii)* What is the optimal level of data compression to be used? *(iii)* What is the amount of data that should be sent through each selected RAN after compression? While addressing the above issues, we account for both network characteristics and application requirements, providing a solution, which achieves an optimal energy-quality-cost tradeoff.

The rest of the paper is organized as follows. Section II discusses the related work and highlights our novel contributions, while Section III introduces the system model under study. Section IV introduces our problem formulation and presents the analytical solution of the problem. Section V describes our iterative algorithm that converges to the optimal solution even when PDAs do not have a reliable initial estimate of their resource share over the available RANs. Section VI presents the performance evaluation of our solution and the gain that it provides with respect to state-of-the-art techniques. Finally, Section VII concludes the paper.

II. Related Work

The investigated approaches in the field of m-health can be broadly classified, according to the overall architecture of m-health systems in Figure 1, into three categories: data acquisition and in-network processing (including low-power hardware designs, smartphone health-related applications, signal compression, and feature extraction [5]), wireless transmission resource allocation and optimization, and data retrieval at the cloud (including data delivery, reconstruction, and classification).

In m-health systems, sensor technologies have been combined with mobile communications in order to enable data collection from heterogeneous sensing units, while providing ubiquitous access

to the data. In this context, PDAs (i.e., smartphones) were initially used both as a sensor and data aggregator [6], where they gather the medical data from BASN and embedded sensors in smartphones, then forward the aggregate traffic to the MHC. In [7], the *processing* capability of smartphones has been considered. The authors described a smartphone-centric architecture, where smartphones are exploited not only as data aggregators but also as sensing, processing, and transmitting devices[1]. However, they consider wireless network as a transport network that can be accessed through either 3G/4G network or through WiFi interfaces. The aforementioned work has not considered the multihoming feature of smartphones to optimally select the best communication interfaces for transmission.

At the MHC, data retrieval, feature extraction, classification, and further analysis can be performed on the received data, in order to accurately evaluate the state of the patient. The conventional cloud computing architecture enables smart devices (e.g., sensors, smartphones) to exchange information with the cloud in order to provide concise and scalable processing as well as storage services for supporting application requirements [9]. However, the deployment of health monitoring and emergency response applications on the cloud are facing challenges due to the unpredictable delay caused by transferring data to and from the cloud. A promising solution to overcome this issue consists in performing efficient in-network processing at the PDA level (i.e., edge user) [10]. This could significantly help to save network resources, offload core network traffic, and meet application requirements for swift and secure data transfer.

Regarding wireless transmission optimization, in future 5G networks, user association with network infrastructure may be concurrent, exploiting the multihoming feature of mobile devices, or to be switched from one access point to another (within the same Radio Access Network (RAN), or between different RANs) in order to enhance system performance and user experience [11]. This switching and concurrent association between different RANs needs to be done while taking into consideration different network characteristics and application requirements. In this context, several initiatives have recently explored the interaction between cellular operators and WiFi network owners [12]. The WLAN community has also participated in this trend with some initiatives, such as Hot Spot 2.0 and high-efficiency WLAN standardization, which study RAN-based integration solutions [13]. These initiatives aim to increase cooperation between 3GPP Long Term Evolution (LTE) and WiFi radio technologies. For instance, the authors in [13] discuss convergence of WLAN-based small cells with operator-managed cellular deployment, and envision different architectural options for networks integration. Also, new standards like IEEE 802.21 exploits interoperability between heterogeneous networks for handover optimization [14]. However, this standard only facilitates

[1]A comprehensive overview of recent smartphone applications designed for remote health monitoring can be found in [8].

handover without specifying any network selection mechanisms[2].

The problem of determining the optimal association between users and available RANs across multiple RATs, operating at different frequencies and using different protocols, has received a great deal of attention. Two main approaches have been proposed: a game theoretic approach, where convergence to Nash equilibrium and the Pareto-efficiency of this equilibrium have been studied [15][16], and a utility-based approach, where decisions are made based on properly defined utility functions and the network with the best value is selected. The utility can be a function of monetary cost, power consumption, network conditions, or user preferences [17][18].

Other studies have considered the resource allocation problem for parallel transmission using multiple RATs [19][20]. However, the formulated problem in [19] is NP-hard, and a sub-optimal allocation strategy is developed by exploiting the intrinsic quasi-concavity of the problem. While, in [20], the authors present a framework of multi-RAT systems, where a small cell serves a number of mobile users via IEEE 802.11 WLAN and 3GPP LTE access technologies. A scheduler at the small cell is proposed to minimize the total transmission power subject to quality of service constraints on the users transmission rates. In [21], an urban deployment scenario is investigated, where WiFi small cells are overlaid on top of the 3GPP LTE network. The authors propose user-centric network selection algorithms to minimize feedback overhead while taking into account user preferences. A comprehensive review of the state-of-the-art mathematical models that are applied to the network selection problem, including utility-based approach, game theory, combinatorial optimization, Markov decision processes, and fuzzy logic can be found in [22].

To the best of our knowledge, none of the aforementioned work advocates the user-centric approach for efficient network association with active context-aware in-network processing in order to improve the delivery, cost, and latency of the patient data to MHC. Our main contributions can be summarized as follows:

1) We formulate a multi-objective problem that allows each PDA to optimally set its data compression ratio and select the RAN(s) for data transmission, in an energy-efficient and cost-aware manner while ensuring an acceptable signal distortion.
2) We propose an analytical solution for the optimization problem, by decomposing it into two sub-optimizations. The two sub-problems turn out to be solvable with low complexity, and they are proved to lead to the same optimal solution as the original problem.
3) We design a distributed, iterative, PDA-centric algorithm and we analytically show that it can converge to the optimal solution starting from any rough estimate of their resource share on the available RANs.

[2]We remark that handover execution, although being a relevant aspect, is not addressed in our work.

4) Finally, we evaluate the performance of the proposed solution and compare it against that of state-of-the-art techniques. Results show that the proposed approach allows for high-quality healthcare monitoring of patients, it significantly outperforms other solutions, and swiftly adapts to varying network conditions.

III. System Model and Performance Metrics

Consider a wireless heterogeneous m-health system, as shown in Figure 2. It is assumed that the medical data[3] is collected and sent to the PDA (i.e., smartphone) which compresses the data using adaptive compression techniques. Each PDA has multiple RANs available, through which it can transmit its data toward the MHC. The RANs have different characteristics, such as data rate, energy consumption, monetary cost (i.e., requested payment for using network services), and transmission delay. Furthermore, due to mobility, wireless channel dynamics and time-varying traffic patterns, the level of quality of service offered by the available RANs may change over time.

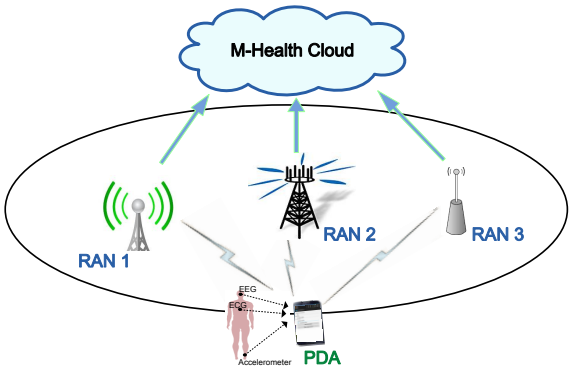

Fig. 2. System scenario under study.

In the following, we focus on a time period T and assume that each PDA i ($i = 1, \dots N$) has to transfer B_i bits of data toward the MHC. As mentioned, data is compressed by each PDA so that the actual amount of bits to be transmitted is given by: $b_i = B_i(1 - \kappa_i)$ with κ_i being the data compression ratio adopted by PDA i. Data compression introduces a signal distortion, which can be expressed through the Percentage Root-mean-square Difference (PRD) between the recovered EEG data and the original one. Using the results obtained through our real-time implementation [24], the relation between encoder distortion D_i, compression ratio κ_i and wavelet filter length F is defined as

$$D_i = \frac{c_1 e^{(1-\kappa_i)} + c_2 \cdot (1 - \kappa_i)^{-c_3} + c_4 \cdot F^{-c_5} - c_6}{100}. \quad (1)$$

[3]Although the proposed framework is assumed to employ the encoding model of EEG signals [23], it can be extended to diverse biosignals and multimedia data, which are typically encoded at higher data rates, such as streaming of electrocardiogram (ECG) signals and medical video, or at a lower data rate, such as human pressure or heart-rate reading.

where κ_i is evaluated as $\kappa_i = 1 - \frac{Q}{S}$, with Q being the number of output samples generated after discrete wavelet transform (DWT)-based EEG compression and S being the length of the input signal, while the model parameters c_1, c_2, c_3, c_4, c_5 and c_6 are estimated by the statistics of the typical EEG encoder used in [24].

Now, assume that RAN j operates on a band of width W_j and that the generic PDA i can enjoy a data rate r_{ij} on RAN j. Clearly, r_{ij} depends on the network access technology (e.g., its maximum value is 54Mbps in IEEE 802.11a/g) and on channel radio propagation conditions. As noted in [23], the estimated energy consumption for PDA i to send b_i bits over RAN j is:

$$\widetilde{E}_{ij} = \psi_j \left(\frac{b_i N_0 W_j}{r_{ij} g_{ij}} (2^{\frac{r_{ij}}{W_j}} - 1) \right) + c_j \,. \tag{2}$$

In the above expression, N_0 is the noise spectral density, while the channel gain g_{ij} is defined as

$$g_{ij} = K \cdot \sigma \cdot |h_{ij}|^2$$

where $K = -1.5/(\log(5\text{BER}))$, σ is the path loss attenuation, and $|h_{ij}|$ is the fading channel magnitude for PDA i over RAN j. In (2), ψ_j and c_j are specific parameters that differ for each network interface [25]. They can be found in the radio interface specifications, or obtained through power consumption measurements [26].

Next, looking at the expected latency provided by each RAN, we define:

$$\widetilde{L}_{ij} = \frac{b_i}{r_{ij}} + \xi_j, \tag{3}$$

where $\frac{b_i}{r_{ij}}$ and ξ_j are, respectively, the air time and the access channel delay that PDA i expects to experience when transmitting b_i bits through RAN j. In other words, it represents the estimated end-to-end delay when using a given technology [27].

Finally, the monetary cost (hereinafter referred to as cost for brevity) resulting from using RAN j by PDA i to send b_i bits is expressed in Euro and defined as:

$$\widetilde{C}_{ij} = b_i \varepsilon_j \tag{4}$$

where ε_j is the monetary cost per bit for RAN j. This monetary cost can be acquired through the use of, e.g., the IEEE 802.21 standard [14], which allows a user device to gather information about the available wireless networks [26]. Such value can also be stored on the PDAs in advance and updated if there are any changes in pricing.

IV. Joint Network Selection and Compression Optimization: Problem Formulation and Solution

Looking at the system model and the performance provided by each RAN (1)–(4), it can be seen that there is a tradeoff between distortion on one side and energy consumption, latency and cost on

the other. First, the higher the compression ratio (κ_i), the greater the distortion, but the smaller the amount of data to transmit (b_i). Secondly, as the data rate over RAN j (r_{ij}) increases, the energy consumption increases, while the latency decreases. Also, it is often the case that RANs providing higher data rates and lower latency have a higher monetary cost.

In a system supporting healthcare applications, it is of paramount importance to provide an acceptable distortion level and to ensure a swift transfer of medical data towards the MHC. Then, from a practical point of view, it is crucial that PDAs do not have to be recharged too often and that services have an acceptable cost. Thus, in light of such requirements, in Section IV-A we formulate a Multi-objective Optimization Problem (MOP) that each PDA should solve and whose aim is to find the optimal tradeoff between the above conflicting objectives. The proposed problem is then analytically solved in Section IV-B.

A. Problem Formulation

The objective of the proposed MOP is threefold: (i) minimizing transmission energy consumption, (ii) minimizing monetary cost, and (iii) meeting the medical data QoS requirements in terms of signal distortion and data delivery latency. We therefore define a single aggregate objective function which turns the above multiple objectives into a single objective function. However, each objective presents different ranges and units of measurement, hence we first normalize these quantities with respect to their maximum value, in order to make them adimensional and comparable. We will denote the normalized energy, monetary cost, and latency by E_{ij}, C_{ij} and L_{ij}, respectively.

Given a generic PDA i with M available RANs, the objective of our optimization problem is to obtain the optimal compression ratio and assign the PDA to the optimal RAN(s) minimizing the transmission energy consumption E_{ij}, monetary cost C_{ij}, latency L_{ij}, and distortion D_i:

$$\textbf{P:} \quad \min_{P_{ij},\kappa_i} \sum_{j=1}^{M} P_{ij}U_{ij} + \delta_i D_i \tag{5}$$

$$\text{s.t.} \quad \frac{P_{ij} \cdot b_i}{r_{ij}} \leq T_{ij}, \quad \forall j \in M \tag{6}$$

$$\sum_{j=1}^{M} P_{ij} \geq 1, \tag{7}$$

$$0 \leq P_{ij} \leq 1, \quad \forall j \in M \tag{8}$$

$$0 \leq \kappa_i \leq 1. \tag{9}$$

where $U_{ij} = \alpha_i E_{ij} + \beta_i C_{ij} + \gamma_i L_{ij}$ is the utility function of PDA i over RAN j. The weighting coefficients represent the relative importance of the four objective functions in the problem; it is

assumed that $\alpha_i + \beta_i + \gamma_i + \delta_i = 1$. Moreover, in (5) we consider a network utilization indicator P_{ij} that represents the fraction of data that should be transmitted through RAN j by PDA i. Note that PDAs have all information to compute the expected energy consumption and cost. Additionally, we assume that the RAN notifies the PDA about the physical data rate r_{ij}. The network also notifies PDAs about the expected channel access delay (ξ_j) the PDA may experience.

The network capacity constraint is represented by (6), where T_{ij} is the maximum fraction of the time period T that can be used by PDA i over RAN j (resource share). T_{ij} depends on the number of PDAs accessing the RAN, and we assume that it is notified by the RAN. Constraint (7) instead ensures that all the data that PDA i has to transfer to the MHC is actually sent through the wireless medium.

The unknowns in this problem are the P_{ij}'s and κ_i, i.e., each PDA needs to determine its compression ratio and the amount of data that the PDA should transfer through the different RANs. Looking at problem formulation in (5), one can see that it is not a linear programming (LP) problem [28], due to the terms involving the product of P_{ij} by κ_i (or functions of κ_i). Thus, below we envision a methodology to decompose[4] the problem into two sub-optimization problems, for which an optimal, analytical solution can be obtained.

B. Optimization Decomposition

In order to analytically solve (5), one would like to break the original problem into two sub-problems such that each of them is a function of one decision variable only and, hence, can be solved independently of the other. The difficult point in our case is that the optimization variables (i.e., P_{ij}'s and κ_i) are coupled. To overcome this issue, we proceed as follows.

We first look at the optimization variables in (5) as network selection variables P_{ij}'s and adaptive compression variables κ_i. Network selection variables can be considered as global variables that are relevant to the overall system, while adaptive compression variables are local variables at each PDA. We therefore decompose the problem into the network selection and adaptive compression sub-problems, and we prove that solving the new problem formulation still leads to the optimal solution of the original problem in (5).

Theorem 1**:** *The optimization problem in (5) can be decomposed into two sub-optimization*

[4]Note that a simple approach, like transforming the problem into a Geometric Program (GP), would not work in this case due to the existence of the constraint in (7) with the non-linearity of the distortion objective.

problems while maintaining the optimal solution, as follows:

$$\textbf{SP1:} \quad \min_{P_{ij}} \sum_{j=1}^{M} P_{ij}\hat{U}_{ij} \tag{10}$$

$$\text{s.t.} \quad \frac{P_{ij}b_i}{r_{ij}} \leq T_{ij}, \quad \forall j \in M \tag{11}$$

$$\sum_{j=1}^{M} P_{ij} \geq 1, \tag{12}$$

$$0 \leq P_{ij} \leq 1, \quad \forall j \in M \tag{13}$$

and

$$\textbf{SP2:} \quad \min_{\kappa_i} \left(\delta_i D_i - \sum_{j=1}^{M} P_{ij}\kappa_i \bar{U}_{ij} \right) \tag{14}$$

$$0 \leq \kappa_i \leq 1 \tag{15}$$

where $\bar{U}_{ij}$ and $\hat{U}_{ij}$ have a similar expression as U_{ij} but for some constant terms and the fact that b_i is replaced by B_i, i.e., they are independent of κ_i.

Proof: See Appendix A. ■

It is worth mentioning that decoupling the overall optimization problem into two sub-problems, greatly simplifies the problem, thus allowing the study of different adaptive compression techniques with different distortion models. Similarly, the usage of various video coding schemes for medical video content delivery could be investigated, in the presence of different network types and network conditions.

C. Network Selection Optimization

In the following, we present an analytical solution for the network selection optimization problem in (10). The problem is an LP problem. Thus, we can reduce the objective function by increasing P_{ij}'s with minimum $\hat{U}_{ij}$. Since these variables have a nonnegative coefficients in z, there would be no other way to decrease z. We conclude that in the following proposition.

Proposition 1**:** *The optimal solution of (10) can be obtained by maximizing the values of P_{ij}'s for which the corresponding $\hat{U}_{ij}$'s are minimum.*

Proof: See Appendix B. ■

According to the above proposition, we can solve the network selection optimization problem in (10) using Algorithm 1, while maintaining the optimal solution. The algorithm sorts the available RANs in ascending order, according to the values $\hat{U}_{ij}$'s, then the network with the lowest $\hat{U}_{ij}$ is selected, and P_{ij} is calculated as:

$$P_{ij} = \max\left(1, \frac{T_{ij}r_{ij}}{b_i}\right). \tag{16}$$

Note, indeed that P_{ij} cannot exceed 1, according to the constraint in (8). If the available resource share on this RAN does not satisfy the requirement of the PDA (i.e., $\sum_{j=1}^{M} P_{ij} < 1$), the second RAN in the list is selected, and so on till the constraint in (7) is satisfied. The algorithm complexity is closely related to the number of available RANs M; the worst-case complexity is $O(M \log(M))$.

Algorithm 1 Network Selection Optimization at PDA i

Require: α_i, β_i, γ_i, T_{ij}, r_{ij}, ε_j

```
p = 1.
Sort the available RANs according to the Û_ij values in ascending order
for j = 1 → M do
    Compute P_ij according to (16)
    if P_ij ≥ p then
        Set P_ij = p, and P_ik = 0, ∀k s.t. j < k ≤ M.
        Break % Constraint (7) is met
    else
        p = p − P_ij
    end if
end for
return Selected RAN(s) and corresponding optimal P_ij's
```

D. Adaptive Compression Optimization

As far as the problem in (14) is concerned, a closed-form expression for the solution can be obtained by imposing that the derivative with respect to κ_i of the objective function is equal to 0. I.e.,

$$\begin{aligned} \partial/\partial\kappa_i &= \partial/\partial\kappa_i \left[\delta_i D_i - \sum_{j=1}^{M} P_{ij}\kappa_i \bar{U}_{ij} \right] \\ &= \delta_i \frac{\partial D_i}{\partial \kappa_i} - \sum_{j=1}^{M} P_{ij}\bar{U}_{ij} = 0. \end{aligned} \tag{17}$$

The distortion in (1) can be approximated[5] as,

$$D_i \approx \frac{c_2(1-\kappa_i)^{-c_3} + c_4 F^{-c_5} - 6.5}{100}. \tag{18}$$

[5]This simplified expression is still extremely accurate with a mean square error that equals 0.1%, while enabling us to maintain a closed-form expression of the solution.

By substituting (18) in (17), we obtain:

$$\begin{aligned} \frac{\partial D_i}{\partial \kappa_i} &= \frac{\sum_{j=1}^{M} P_{ij}\bar{U}_{ij}}{\delta_i} \\ \frac{c_2 c_3 (1-\kappa_i)^{-c_3-1}}{100} &= \frac{\sum_{j=1}^{M} P_{ij}\cdot\bar{U}_{ij}}{\delta_i} \\ (1-\kappa_i)^{-c_3-1} &= \frac{100\sum_{j=1}^{M} P_{ij}\cdot\bar{U}_{ij}}{\delta_i\cdot c_2\cdot c_3} \\ \log(1-\kappa_i) &= -\frac{\log\zeta}{(1+c_3)} \end{aligned} \tag{19}$$

where

$$\zeta = \frac{100\sum_{j=1}^{M} P_{ij}\cdot\bar{U}_{ij}}{\delta_i c_2 c_3}.$$

Thus, according to (19), the optimal κ_i is given by:

$$\kappa_i = 1 - \zeta^{-\frac{1}{1+c_3}}\,. \tag{20}$$

V. Adaptive Network Selection and Compression

In this section, we propose a distributed, iterative algorithm for optimal Adaptive Network Selection and Compression, named ANSC for short. ANSC leverages the problem decomposition introduced in the previous section, and it aims at finding the optimal solution of (5) in practical scenarios where PDAs may have just an initial estimate of their resource share on a given RAN j.

According to ANSC, once obtained the list of the available RANs, each PDA i initially assumes that no compression is performed (i.e., $\kappa_i = 0$) and runs Algorithm 1 locally, in order to find the optimal values of P_{ij} that determine which network(s) i should use and the amount of data that i should transfer on each RAN. Recall that in Algorithm 1 the weights α_i, β_i and γ_i are assumed to be pre-defined according to application requirements and/or PDAs' preferences. Importantly, the value T_{ij} is initially set to $T_{ij} = T_j/N_j, \forall j$, where N_j is the number of PDAs using RAN j, i.e., all PDAs assume to receive the same resource share on RAN j [6]. As foreseen by several standards, the RAN can notify users about the value of N_j.

Next, the generic PDA obtains the optimal κ_i using Eq. (20). It then broadcasts the corresponding value of $\widetilde{T}_{ij}$'s, i.e., the amount of resources it intends to "consume" over RAN j, which is given by:

$$\widetilde{T}_{ij} = \frac{P_{ij} b_i}{r_{ij}}\,. \tag{21}$$

At the RAN point of access, the actual demand from all PDAs is calculated, and each RAN j can use whatever mechanism to allocate the remaining resources among competing users (e.g., using

[6]Note that the value of T_{ij} can be initially set to any arbitrary value.

proportional fairness, round robin, etc.) [29]. Each RAN can then return to the PDAs the values of their actual resource share. Accordingly, the PDAs run network selection optimization (Alg. 1) again, obtaining the updated optimal values of P_{ij}'s. The procedure can be repeated until convergence or a maximum number of iterations have been reached. The main steps of the ANSC algorithm are illustrated in Algorithm 2.

Algorithm 2 Adaptive Network Selection and Compression (ANSC) algorithm at PDA i

Initialization: $\kappa_i = 0$, $t = 0$
$S(t) = 0$
do
Get optimal P_{ij}'s through Algorithm 1
Compute κ_i using (20)
Broadcast requested $\widetilde{T}_{ij}$'s computed through (21)
Get updated $T_{ij}(t+1)$ from RAN j
$t++$
$S(t) \leftarrow$ Value of objective function in (5)
while $|S(t) - S(t-1)| > \epsilon \wedge t < n_{iter}$
return Selected RAN(s); optimal P_{ij}'s and κ_i

Below, we prove that the convergence of ANSC scheme is guaranteed.

Theorem 2**:** *Regardless of the scheduling mechanisms implemented at the available RANs, the ANSC scheme converges to the optimal solution of the optimization problem in (5).*

Proof: See Appendix C. ■

We remark that the algorithm naturally converges when $\widetilde{T}_{ij}(t+1) = \widetilde{T}_{ij}(t)$, i.e., when the PDAs are not willing to give away any fraction of their resource share on the RANs. However, due to network dynamics, the available resource shares on the RANs as well as the PDAs traffic demand may vary: this may trigger the PDAs to run the ANSC algorithm again and update their resource allocation.

VI. Simulation Results

We now evaluate the system performance and show the convergence behavior via simulation. We compare the performance of the proposed ANSC scheme against two baseline algorithms: the Enhanced Power-Friendly Access Network Selection (EPoFANS) algorithm, which implements the scheme presented in [26], and the Autonomous Access Network Selection (AANS) algorithm presented in [30]. To this end, we consider the network topology shown in Figure 2, where each

PDA can connect to four RANs with different characteristics. Specifically, RAN_1 with a monetary cost per bit $\varepsilon_1 = 610^{-6}$ Euro/bit and data rate $r_1 = 4$ Mbps; RAN_2 with $\varepsilon_2 = 310^{-6}$ Euro/bit and $r_2 = 2.5$ Mbps, RAN_3 with $\varepsilon_3 = 0$ Euro/bit, $r_3 = 1.5$ Mbps; RAN_4 with $\varepsilon_4 = 110^{-6}$ Euro/bit and $r_4 = 2$ Mbps. Moreover, to emphasize the tradeoff between distortion, energy consumption, latency and cost, it is assumed that $\xi_j = 0, \forall j \in M$.

A PDA can capture 4096 samples of epileptic EEG data [31]. Each raw sample is represented using 12 bits. As far as the channel dynamics are concerned, flat Rayleigh fading is assumed, with Doppler frequency of 0.1 Hz. The other physical layer parameters over the available RANs are set to: noise spectrum density $N_0 = -174$ dBm, bandwidth $W = 0.5$ MHz, and path loss attenuation $\sigma = 3.6 * 10^{-6}$.

First, in order to assess the importance of optimizing both network selection and the compression ratio, Figure 3 depicts the value of the objective function in (5) as the compression ratio κ_i varies, when $\alpha_i = \beta_i = \gamma_i = \delta_i = 0.25$. One can clearly see that with increasing κ_i, the length of the transmitted data decreases; hence, initially the value of the objective function decreases as well. However, beyond a certain value, distortion becomes dominant and the value of the objective function starts increasing. Using a high compression ratio enables PDAs to decrease their load on costly networks and stick to low-cost networks, as shown by Figure 3-(b). This further confirms that, in order to optimize performance, it is important to jointly consider network selection and adaptive compression.

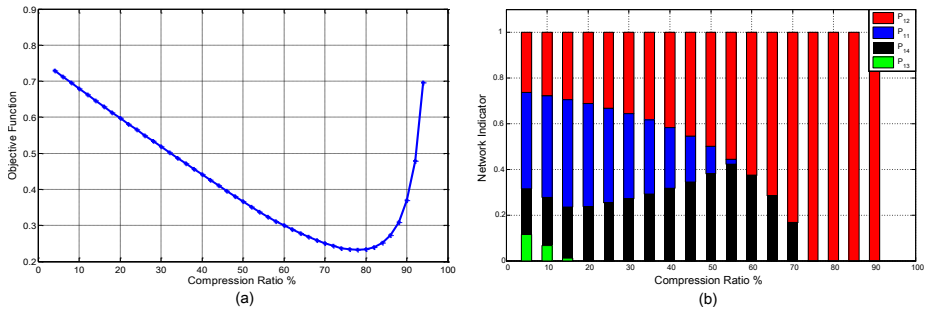

Fig. 3. Value of the objective function (a), and of P_{ij} (network indicators), as the compression ratio, hence distortion, varies.

Next, we compare the performance of the proposed ANSC scheme against the two baseline algorithms. The EPoFANS algorithm computes a score for each of the candidate RANs, using the same utility function as U_{ij}. It then selects the network with the lowest score as a target network [26]. For sake of fairness, here we enhance EPoFANS with the adaptive compression optimization in (14) to obtain the optimal value of compression ratio κ_i. In the AANS algorithm, instead, we fix κ_i to a certain value, and determine the optimal RAN(s) by solving the optimization problem in (10) [30]. Furthermore, we assess the ability of the tested schemes to adapt to network dynamics, in

particular, we assume that the number of PDAs that can access the available RANs varies over time, as shown in Figure 4-(b). As expected, Figure 4-(a) shows that, when the number of PDAs decreases, the resource share for a generic PDA grows and the value of the aggregate objective functions drops for all schemes. The opposite occurs when new PDAs join the network. Interestingly, the network quickly adapts to any change in the scenario by assigning more or less resources to the PDAs and swiftly reaching convergence to the optimum. In all cases, however, ANSC provides the best performance.

Figure 5 presents the value of different performance metrics when ANSC, AANS and E-PoFANS are adopted, and the values of the corresponding network indicators P_{ij}'s over the different RANs. Here, we fixed the resource share available to each PDA to be $T_{ij} = T_j/N_j \ \forall j$. We remark that E-PoFANS selects only one network (the one with the lowest score), thus in this case P_{ij} takes a value equal to either 0 or 1 (see Figure 5-(H)). Our scheme and AANS instead take different candidate networks into account and select the optimal RAN(s) that minimize the PDA's aggregate objective, i.e., P_{ij} can take any value between 0 and 1. It follows that PDAs can transmit using different RANs simultaneously instead of being limited to one RAN only (see Figure 5-(E)). This enables PDAs to decrease their load on costly networks and distribute it on low-cost networks, which results in a reduced energy consumption and monetary cost (see Figure 5-(a),(b)).

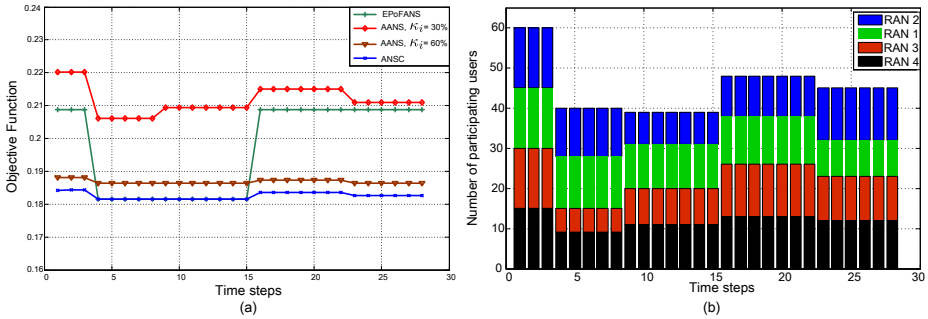

Fig. 4. Temporal evolution of the system performance: aggregate objective function (a) as the number of PDAs varies.

Importantly, unlike AANS that uses a fixed compression ratio, our scheme finds the compression ratio corresponding to the optimal tradeoff between different performance metrics. Herein, we consider two possible values of compression ratio for AANS: a low κ_i, namely 30%, and a high κ_i, namely 60%. The former results to be lower than the optimal compression ratio obtained with ANSC, thus it leads to more transmitted bits. As a consequence, AANS gives a higher energy consumption and and monetary cost (see Figure 5-(a),(b)), as it increases the amount of data transmitted on costly networks (see Figure 5-(F)). On the contrary, the latter value ($\kappa_i = 60\%$) is higher than the optimum. Despite the decrease in transmission energy, monetary cost, and latency due to the smaller amount

of transmitted data (see Figure 5-(a),(b),(C)), AANS leads to a higher objective function because of the large distortion (see Figure 5-(d)). Thus, from Figure 4 and Figure 5, we can conclude that our ANSC scheme leads to the optimal tradeoff among the target performance metrics, while other presented algorithms focus on one or more performance metric at the expense of the others.

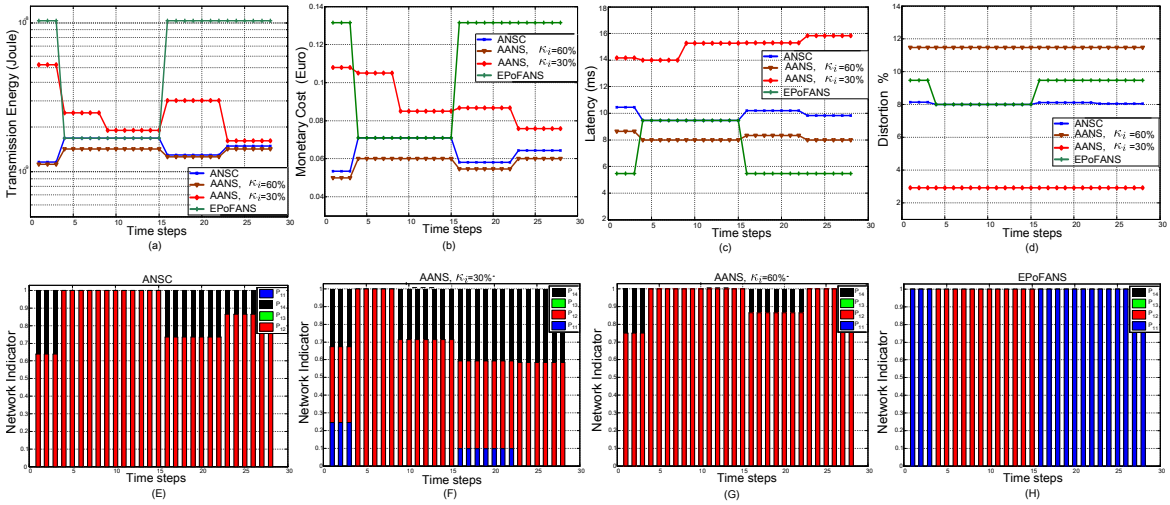

Fig. 5. A comparison of different performance metrics and network indicators under ANSC, AANS and E-PoFANS, with varying number of PDAs.

Finally, Figure 6 depicts the convergence behavior of the ANSC scheme, compared to AANS. In this case, we combine AANS with exhaustive search (AANS-ES) so as to iteratively solve the optimization problem in (5) and find the optimal P_{ij}'s for each κ_i. Specifically, in AANS-ES, initially $\kappa_i = 0$, then it is incremented by a small quantity at every iteration. On the contrary, recall that ANSC leverages the problem decomposition into two sub-optimization problems. Although the mechanism exploited by ANSC is iterative, we observe that very few iterations are needed in order to reach convergence, compared with AANS-ES.

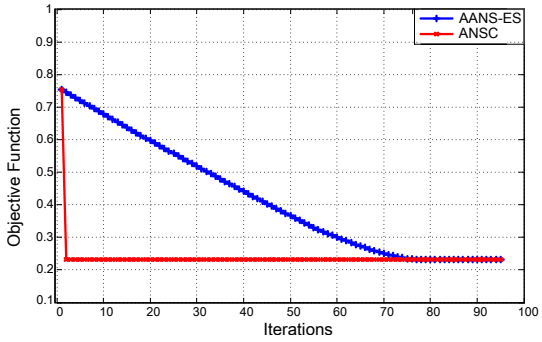

Fig. 6. Convergence behavior of the proposed ANSC scheme and of AANS with exhaustive search.

VII. Conclusions

We considered a heterogeneous m-health system where multiple radio access technologies can be simultaneously exploited by PDAs, in order to improve the performance of data transfer over the wireless medium. We proposed a dynamic network selection mechanism that enables energy efficient and high quality patient health monitoring by targeting jointly RAN selection and data compression. In the proposed scheme, energy consumption, the application quality of service requirements, and monetary cost are considered as main performance metrics and integrated into a multi-objective optimization problem. We proved that the optimal solution to the problem can be obtained analytically by decomposing the problem into two sub-problems. The two sub-problems have low complexity and allow for a swift solution at the PDAs. Simulation results show that our scheme significantly outperforms existing techniques, as well as its ability to adapt to varying network conditions.

References

[1] A. Solanas, et al., "Smart health: A context-aware health paradigm within smart cities," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 74–81, Aug 2014.

[2] N. Roy, C. Julien, A. Misra, and S. K. Das, "Quality and context-aware smart health care: Evaluating the cost-quality dynamics," *IEEE Systems, Man, and Cybernetics Magazine*, vol. 2, no. 2, pp. 15–25, April 2016.

[3] M. Yuce, S. Ng, N. Myo, J. Khan, and W. Liu, "Wireless body sensor network using medical implant band," *J. Medical Systems*, vol. 31, no. 6, pp. 467–474, 2007.

[4] J. Andrews, S. Singh, Q. Ye, X. Lin, and H. Dhillon, "An overview of load balancing in hetnets: old myths and open problems," *IEEE Wireless Communications*, vol. 21, no. 2, pp. 18–25, April 2014.

[5] J. Chiang and R. K. Ward, "Energy-efficient data reduction techniques for wireless seizure detection systems," *Sensors 2014*, pp. 2036–2051, 2014.

[6] S. Adibi, "Link technologies and blackberry mobile health (mhealth) solutions: A review," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 4, pp. 586–597, July 2012.

[7] I. Bisio, F. Lavagetto, M. Marchese, and A. Sciarrone, "A smartphone-centric platform for remote health monitoring of heart failure," *Wiley International Journal of Communication Systems*, vol. 28, no. 11, pp. 1753–1771, July 2015.

[8] J. Wang et al., "Smartphone interventions for long-term health management of chronic diseases: An integrative review," *Telemedicine and e-Health*, vol. 20, no. 6, pp. 570–583, June 2014.

[9] S. Barbarossa, S. Sardellitti, and P. D. Lorenzo, "Communicating while computing: Distributed mobile cloud computing over 5g heterogeneous networks," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 45–55, Nov 2014.

[10] I. Bisio, F. Lavagetto, M. Marchese, and A. Sciarrone, "Smartphone-centric ambient assisted living platform for patients suffering from co-morbidities monitoring," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 34–41, January 2015.

[11] N. Himayat, S.-P. Yeh, A. Panah, S. Talwar, M. Gerasimenko, S. Andreev, and Y. Koucheryavy, "Multi-radio heterogeneous networks: Architectures and performance," in *International Conference on Computing, Networking and Communications (ICNC)*, Feb 2014, pp. 252–258.

[12] L. Gao, G. Iosidis, J. Huang, L. Tassiulas, and D. Li, "Bargaining-based mobile data ofoading," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.

[13] S. Andreev, M. Gerasimenko, O. Galinina, Y. Koucheryavy, N. Himayat, S.-P. Yeh, and S. Talwar, "Intelligent access network selection in converged multi-radio heterogeneous networks," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 86–96, December 2014.

[14] "IEEE Standard for Local and Metropolitan Area Networks−Part 21: Media Independent Handover," *IEEE Std. 802.21-2008*, 2009.

[15] E. Aryafar, M. W. A. Keshavarz-Haddad, and M. Chiang, "RAT selection games in HetNets," *IEEE INFOCOM*, pp. 998–1006, Apr 2013.

[16] K. Zhu, E. Hossain, and D. Niyato, "Pricing, spectrum sharing, and service selection in two-tier small cell networks: A hierarchical dynamic game approach," *IEEE Transactions on Mobile Computing*, vol. 13, no. 8, Aug. 2014.

[17] H. Liu, C. Maciocco, V. Kesavan, and A. Low, "Energy efficient network selection and seamless handovers in mixed networks," in *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks Workshops*, June 2009, pp. 1–9.

[18] H. Petander, "Energy-aware network selection using traffic estimation," *ACM workshop on Mobile internet through cellular networks*, pp. 55–60, 2009.

[19] G. Lim, C. Xiong, L. Cimini, and G. Li, "Energy-efficient resource allocation for ofdma-based multi-rat networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2696–2705, May 2014.

[20] M. Fadel, A. Ibrahim, and H. Elgebaly, "Qos-aware multi-rat resource allocation with minimum transmit power in multiuser ofdm system," in *IEEE Globecom Workshops (GC Wkshps)*, Dec 2012, pp. 670–675.

[21] M. Gerasimenko, N. Himayat, S.-P. Yeh, S. Talwar, S. Andreev, and Y. Koucheryavy, "Characterizing performance of load-aware network selection in multi-radio (wifi/lte) heterogeneous networks," in *IEEE Globecom Workshops (GC Wkshps)*, Dec 2013, pp. 397–402.

[22] L. Wang and G. Kuo, "Mathematical modeling for network selection in heterogeneous wireless networks-A tutorial," *IEEE Communications Surveys Tutorials*, vol. 15, no. 1, pp. 271–292, 2013.

[23] A. Awad, A. Mohamed, A. A. El-Sherif, and O. A. Nasr, "Interference-aware energy-efficient cross-layer design for healthcare monitoring applications," *Comput. Netw., vol. 74*, pp. 64–77, Dec. 2014.

[24] A. Awad, M. Hamdy, A. Mohamed, and H. Alnuweiri, "Real-time implementation and evaluation of an adaptive energy-aware data compression for wireless EEG monitoring systems," *10th International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (QSHINE)*, pp. 108–114, Aug. 2014.

[25] K. Mahmud, M. Inoue, H. Murakami, M. Hasegawa, and H. Morikawa, "Measurement and usage of power consumption parameters of wireless interfaces in energy-aware multi-service mobile terminals," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 2, pp. 1090–1094, Sept 2004.

[26] R. Trestian, O. Ormond, and G.-M. Muntean, "Enhanced power-friendly access network selection strategy for multimedia delivery over heterogeneous wireless networks," *IEEE TRANSACTIONS ON BROADCASTING, vol.60, no.1*, pp. 85–101, Mar. 2014.

[27] Y. Wang, M. C. Vuran, and S. Goddard, "Cross-layer analysis of the end-to-end delay distribution in wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 20, no. 1, pp. 305–318, Feb 2012.

[28] S. Boyd and L. Vandenberghe, *Convex Optimization*, 1st ed. cambridge university press, 2003.

[29] M. Salem, A. Adinoyi, M. Rahman, H. Yanikomeroglu, D. Falconer, Y.-D. Kim, E. Kim, and Y.-C. Cheong, "An overview of radio resource management in relay-enhanced OFDMA-based networks," *IEEE Communications Surveys Tutorials*, vol. 12, no. 3, pp. 422–438, Third 2010.

[30] A. Awad, A. Mohamed, and C. F. Chiasserini, "User-centric network selection in Multi-RAT systems," *IEEE WCNC workshop on Mobile Edge Computing and IoT*, April 2016.

[31] R. Andrzejak, K. Lehnertz, C. Rieke, F. Mormann, P. David, and C. Elger, "Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev. E, 64, 061907, (2001)*, 2001.

APPENDIX

A. *Proof of Theorem* 1

The distortion term D_i (see Eq. (1)) is not a function of P_{ij}. Thus, the objective function in (5) can be written as

$$\min_{P_{ij},\kappa_i} \sum_{j=1}^{M} P_{ij}U_{ij} + \min_{\kappa_i} \ \delta_i D_i \,. \tag{22}$$

By denoting with E_M, C_M and L_M, the maximum energy expenditure, cost and latency, respectively, we can rewrite Eqs. (2), (4) and (3), as:

$$\begin{aligned} E_{ij} &= \frac{\left[(1-\kappa_i)\psi_j\left(\frac{B_i N_0 w_j}{r_{ij} g_{ii}}(2^{r_{ij}/w_j}-1)\right)+c_j\right]}{E_M} \\ &= (1-\kappa_i)\bar{E}_{ij} \\ C_{ij} &= \frac{(1-\kappa_i)B_i\varepsilon_j)}{C_M} = (1-\kappa_i)\bar{C}_{ij} \\ L_{ij} &= \frac{(1-\kappa_i)B_i}{r_{ij}L_M} + \frac{\xi_j}{L_M} = (1-\kappa_i)\bar{L}_{ij} + \hat{L}_{ij} \end{aligned}$$

where $\bar{E}_{ij}$, $\bar{C}_{ij}$, $\bar{L}_{ij}$, and $\hat{L}_{ij}$ are independent of κ_i (as well as P_{ij}). By substituting (23) in (22), we get

$$\begin{aligned} Z &= \min_{P_{ij},\kappa_i} \sum_{j=1}^{M} P_{ij}\left(\alpha_i\bar{E}_{ij} + \beta_i\bar{C}_{ij} + \gamma_i(\bar{L}_{ij}+\hat{L}_{ij})\right) \\ &\quad - \sum_{j=1}^{M} P_{ij}\kappa_i(\alpha_i\bar{E}_{ij} + \beta_i\bar{C}_{ij} + \gamma_i\bar{L}_{ij}) \\ &\overset{(a)}{=} \min_{P_{ij}} \sum_{j=1}^{M} P_{ij}\hat{U}_{ij} - \min_{P_{ij},\kappa_i} \sum_{j=1}^{M} P_{ij}\kappa_i\bar{U}_{ij} \\ &= \min_{P_{ij}} \sum_{j=1}^{M} P_{ij}\hat{U}_{ij} + \max_{P_{ij},\kappa_i} \sum_{j=1}^{M} P_{ij}\kappa_i\bar{U}_{ij} \end{aligned} \tag{23}$$

where in *(a)* $\hat{U}_{ij} = \bar{E}_{ij} + \bar{C}_{ij} + (\bar{L}_{ij} + \hat{L}_{ij})$ and $\bar{U}_{ij} = \bar{E}_{ij} + \bar{C}_{ij} + \bar{L}_{ij}$. Now, to minimize Z, we need to minimize $\sum_{j=1}^{M} P_{ij}\hat{U}_{ij}$ and at the same time maximize $\sum_{j=1}^{M} P_{ij}\kappa_i\bar{U}_{ij}$. However, to maximize $\sum_{j=1}^{M} P_{ij}\kappa_i\bar{U}_{ij}$, we need to maximize $P_{ij}\bar{U}_{ij}$, where $\hat{U}_{ij}$ differs from $\bar{U}_{ij}$ by an additive positive constant. Thus, this is conflict with the minimization of $\sum_{j=1}^{M} P_{ij}\bar{U}_{ij}$. Thus, the only possible solution is to minimize $\sum_{j=1}^{M} P_{ij}\hat{U}_{ij}$ with respect to P_{ij} and maximize $\sum_{j=1}^{M} P_{ij}\kappa_i\bar{U}_{ij}$ with respect to κ_i. Accordingly, we will have:

$$\begin{aligned} Z &= \min_{P_{ij}} \sum_{j=1}^{M} P_{ij}\hat{U}_{ij} + \max_{\kappa_i} \sum_{j=1}^{M} P_{ij}\kappa_i\bar{U}_{ij} \\ &= \min_{P_{ij}} \sum_{j=1}^{M} P_{ij}\hat{U}_{ij} - \min_{\kappa_i} \sum_{j=1}^{M} P_{ij}\kappa_i\bar{U}_{ij}. \end{aligned} \tag{24}$$

This proves that the original optimization problem in (5) can be decomposed into two sub-optimization problems, (10) and (14), which still lead to the optimal solution.

B. *Proof of Proposition* 1

Initially, assume that $\mathbf{P}^* = 0$ is the optimal feasible solution, where $\mathbf{P}^*$ is a vector of the variables P_{ij}'s, with corresponding objective value $\mathbf{Z}^* = 0$. However, the values of P_{ij}'s have to be increased in order to satisfy the constraint in (6), i.e., the optimal solution must be $\mathbf{P}^* > 0$. Next, recall that $\hat{U}_{ij}$ depends on the characteristics of RAN j (e.g., data rate, bandwidth, cost per bit, etc.). By contradiction we can show that the optimum solution $\mathbf{P}^*$ is the one for which the values $\hat{U}_{ij}$'s are minimum, i.e., $\mathbf{P}^* = \tilde{\mathbf{P}}$ such that $\tilde{U}_{ij} = \min \hat{U}_{ij}\ \forall j$. Indeed, consider that $\mathbf{P}^* = \tilde{\mathbf{P}} + \Delta$ instead. In this case, $\hat{\mathbf{U}}(P^*) > \hat{\mathbf{U}}(\tilde{\mathbf{P}})$ with $\hat{\mathbf{U}}(\mathbf{P})$ being the vector of the values $\hat{U}_{ij}$'s obtained for $\mathbf{P}$, and $\mathbf{P}^*\hat{\mathbf{U}}(\mathbf{P}^*) > \tilde{\mathbf{P}}\hat{\mathbf{U}}(\tilde{\mathbf{P}})$. This contradicts the assumption that $\tilde{\mathbf{P}} + \Delta$ is the optimal solution, hence the only optimal solution is $\mathbf{P}^* = \tilde{\mathbf{P}}$; any other feasible solution will have a strictly larger objective value.

C. *Proof of Theorem* 2

The ANSC algorithm initially starts assuming an equal resource share among PDAs and no compression $\kappa_i = 0$. Since instead compression can be used, the actual resource share consumed by any PDA will be less than or equal to this initial arbitrary value, as long as no new users join the RANs (i.e., N_j is fixed). In other words, after obtaining the optimal κ_i using (20), the length of the transmitted data, hence the actual amount of radio resources consumed by the PDAs, decreases. As a result, there will be some extra share of resources on certain RANs that were not available at the previous iteration. No matter the scheduling mechanism used at the RANs to reallocate the free resources, for any arbitrary PDA i, we will have:

$$T_{ij}(t+1) \geq T_{ij}(t), \quad \forall t, j. \tag{25}$$

As a result, at each iteration the constraint in (6) becomes looser and looser, and the objective function can be decreased by increasing the value of the P_{ij}'s corresponding to the lowest U_{ij}'s. Thus, from (25) we can conclude that the objective function will always decay as the number of iterations increases until convergence is reached, and this will happen regardless of the scheduling mechanism implemented at the available RANs.